Münster i/W. 1934. No 4 Phil. Diss

Theoretische Physik

Zur axiomatischen Grundlegung der Wahrscheinlichkeitsrechnung.

Inaugural-Dissertation

zur Erlangung der Doktorwürde
der Philosophischen und Naturwissenschaftlichen Fakultät
der Westfälischen Wilhelms-Universität zu Münster i. Westf.

Vorgelegt am 1. Febr. 1933

von

Johannes Blume
aus Paderborn.

Tag der mündlichen Prüfung: 6. 5. 1933.

Dekan: Herr Prof. Dr. Ley.
Referent: Herr Prof. Dr. Kratzer.

1934
Druck: Heinr. Pöppinghaus, o. H.-G., Bochum-Langendreer.

R. UNIV.
BIBLIOTHEEK
LEIDEN.

Meiner lieben Frau Hildegard!

Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41547224_0050

Zur axiomatischen Grundlegung der Wahrscheinlichkeitsrechnung.

§ 1. Darstellung und Kritik der logischen Grundlage der Wahrscheinlichkeitsrechnung von v. Mises.

§ 2. Definition eines Kollektivs endlich vieler Elemente und Durchführung der Grundoperationen.

§ 3. Der erste Fundamentalsatz mit seinen Folgerungen.

§ 4. Das Bayes'sche Problem.

§ 5. Kollektive mit geometrischer Verteilung.

LITERATUR-VERZEICHNIS.

R. v. Mises: Vorlesungen aus dem Gebiete der angewandten Mathematik. Bd. I „Wahrscheinlichkeitsrechnung“ (1931). (In der Arbeit abgekürzt zitiert als „Bd. I“.)
Schriften zur Wissenschaftlichen Weltauffassung Bd. 3 „Wahrscheinlichkeit, Statistik und Wahrheit“.
Naturwissenschaften: 18, 1930. S. 145 ff. Über kausale und statistische Gesetzmäßigkeit in der Physik.“

H. Feigel: Wahrscheinlichkeit und Erfahrung. „Erkenntnis“ Bd. I Heft 2—4, S. 249 ff.

H. Reichenbach: Kausalität und Wahrscheinlichkeit. „Erkenntnis“ Bd. I Heft 2—4, S. 158 ff.

Zilsel: Besprechung der Wahrscheinlichkeitsrechnung von v. Mises. „Naturwissenschaften“ 25, 1932 S. 472 ff.

A. Kratzer: Kollegheft über „Wahrscheinlichkeit“, hergestellt von der Freien Math. Fachschaft der Universität Münster.

„Diskussion über Wahrscheinlichkeit“. „Erkenntnis“. Bd. I Heft 2—4.

Motto: Ist denn im Zufall ein Gesetz?

VORBEMERKUNG.

Wenn man sich mit dem Problem beschäftigt, Wahrscheinlichkeitsvorgänge mathematisch zu erfassen, könnte es scheinen, man wolle sich einer widerspruchsvollen Aufgabe unterziehen. Man fragt sich nämlich, wie es irgend möglich sei, daß eine Disziplin wie die Mathematik, in der die volle Strenge unerbittlicher Logik gilt, auf Vorgänge angewandt werden könne, denen im Prinzip etwas „Unberechenbares" innezuwohnen scheint. Wenn auch dieser Zwiespalt sogar zu dem Versuch der Begründung einer neuen Logik „der Wahrscheinlichkeitslogik" geführt hat, so ist doch zu bedenken, daß dieser Zwiespalt nur dadurch möglich wird, daß man mit dem Wort „Wahrscheinlichkeit" einen Begriff glaubt verbinden zu können oder müssen, der ohne die Zuordnung eines mathematischen Sinnes bestehen könne. Demgegenüber hat v. M i s e s mit allem Nachdruck immer wieder betont, daß nach seiner Überzeugung eine Wahrscheinlichkeitsrechnung in der Methode ihres axiomatischen Aufbaus und ihrer Anwendung auf empirische Dinge sich grundsätzlich nicht von irgendeiner anderen mathematischen Naturwissenschaft unterscheidet. Nun ist allgemein anerkannt, daß die wissenschaftliche Leistung von v. M i s e s in der Wahrscheinlichkeitsproblematik außerordentlich fruchtbringend ist und von dem Ergebnis einer notwendigen kritischen Untersuchung der axiomatischen Fundamentierung seiner Theorie unberührt bleibt. Diese hat allerdings zu starken Bedenken Anlaß gegeben, die u. a. von Z i l s e l und F e i g e l dargelegt worden sind l. c. Interessanterweise befassen sich diese Bedenken vor allem mit dem Umstand, daß der Wahrscheinlichkeitsbegriff bei v. M i s e s keiner rein mathematischen Interpretation zugänglich ist, sodaß sich seine Theorie hierin doch wieder von jeder anderen mathematischen Naturwissenschaft unterscheidet. Dies ist natürlich kein Zufall. Denn auch der v. M i s e s'schen Theorie liegt eine Tendenz zugrunde, die logisch zwiespältig ist, da sie eine Folge von Einzelvorgängen,

deren Ergebnis im Einzelfall als vollkommen unberechenbar angenommen wird, wie z. B. das Ergebnis eines Einzelwurfs beim Würfeln, einer mathematischen Gesetzmäßigkeit unterwerfen will. Feigel l. c. hat diese Dinge ausführlich dargelegt, und es ist für unsere Untersuchung wichtig, kurz zu besprechen, wie von Mises dieses Problem glaubt lösen zu können. Er setzt voraus, daß über die Ergebnisse einer endlichen Anzahl von Einzelvorgängen nichts ausgesagt werden kann, daß aber bei gedachter unbegrenzter Fortsetzung der Folge eine bestimmte mathematische Aussage über den Charakter der Folge möglich sei, d. h. z. B. beim Würfeln, es soll im Einzelnen alles möglich sein, nur soll ein Grenzwert für die relative Häufigkeit eines Merkmals existieren. Der Kalkül seiner Theorie gilt nur für Folgen mit unendlich vielen Einzelgliedern, in der Praxis liegen aber nur finite Folgen vor, sodaß sich die Frage erhob, inwiefern eine durch die Praxis gelieferte Folge als eine solche Annäherung an das Gebilde einer infiniten Folge angesehen werden könnte, daß die Rechnung mit dieser eine Aussagemöglichkeit für jene ergibt. Dieses Problem ist eingehend geprüft, und es hat sich als unausbleiblich herausgestellt, daß ohne die Voraussetzung einer Gesetzmäßigkeit im Bau einer Folge über ihr Verhalten im Unendlichen keine Aussage denkbar ist. Gesetzmäßigkeit verträgt sich aber nicht mit der Annahme, daß im Endlichen grundsätzlich alles möglich sei, und man erkennt, daß auch die Heranziehung des infiniten Charakters für eine Folge den Zwiespalt nicht beheben kann. Nimmt man diese Verhältnisse als gegeben an, so bleibt nichts anderes übrig, als die Vorstellung aufzugeben, daß sich absolute Willkür im Einzelfall mit der Annahme einer Gesetzmäßigkeit für die Gesamtheit dieser Einzelfälle logisch vereinbaren ließe. Dann stehen zwei Wege offen: Entweder behält man die Konvergenz und damit den infiniten Charakter der Folgen bei oder man legt finite Folgen, wie die Praxis sie liefert, zugrunde. Der erste Weg ist mehrfach beschritten worden [1]). Über den zweiten Weg hat sich von Mises wiederholt entschieden ablehnend geäußert und ihn als völlig ungeeignet bezeichnet. (Siehe Bd. I „Erkenntnis" Jg. 1930 S. 279.) Der Verfasser ist Professor Kratzer zum Dank verpflichtet, daß er ihn angeregt hat, dennoch diesen zweiten Weg zu versuchen. Eine nähere Prüfung ergab, daß die finiten Folgen sehr erfolgversprechende mathematische Eigenschaften aufweisen, die man unmöglich

1) Vergl. u. a.: Dörge, Math. Zeitschrift 32 (1930); Kamke, Einführung in die Wahrscheinlichkeitsrechnung (Hirzel, 1932); Iglisch, Zum Aufbau der Wahrscheinlichkeitsrechnung. Math. Analen Bd. 107 Heft 3 (1932).

unbeachtet lassen kann. Für sie gelten bei sinngemäßer Formulierung nicht nur die Grundoperationen, sondern auch die Fundamentalsätze der v. M i s e s'schen Theorie. Darüber hinaus besteht ein Satz über finite Kollektive, der ein plausibles Kriterium für deren Verbindbarkeit liefert und einen Zusammenhang zwischen Verbindbarkeit und Regellosigkeit aufdeckt, der in der v. M i s e s'schen Theorie vermißt wird. Dies alles ermutigt zu dem Versuch, auf einer Rechnung finiter Kollektive eine neue Wahrscheinlichkeitsrechnung aufzubauen.

Die folgende Untersuchung stellt sich drei Aufgaben: Zunächst soll dargetan werden, daß die axiomatische Grundlegung der v. M i s e s'schen Theorie unhaltbar ist. Dann wird ein endliches Kollektiv definiert und gezeigt, daß die Grundoperationen und Fundamentalsätze der von M i s e s'schen Theorie sinngemäß gültig sind. Soweit geht der Inhalt der vorliegenden Arbeit, die damit den gesamten Kalkül der v o n M i s e s'schen Theorie und deren Anwendungsbereich der neuen Theorie sichert. In einer zweiten Arbeit soll dann die Rechnung mit finiten Kollektiven weiter ausgebaut, der Satz über Verbindbarkeit abgeleitet und eine neue Wahrscheinlichkeitsrechnung versucht werden *).

*) Anmerkung während der Korrektur: Die Fortsetzungsarbeit ist inzwischen fertiggestellt. D. V.

§ 1.

Darstellung und Kritik der logischen Grundlegung der Wahrscheinlichkeitsrechnung von v. Mises.

Der Sinn des Aufbaus der Wahrscheinlichkeitsrechnung von v. Mises ist, 1. ein mathematisches Begriffsgebilde, das Kollektiv, und 2. vier mathematische Grundoperationen zu definieren, die es gestatten, aus gegebenen Kollektiven neue abzuleiten und deren Wahrscheinlichkeitsverteilung aus denen der Ausgangskollektive zu berechnen. Um den Begriff „Kollektiv" festzulegen, definiert v. Mises eine Zahlenmenge, für die ein Erzeugungs-, Konvergenz- und Regellosigkeitsprinzip erfüllt ist. Das Erzeugungsprinzip muß den Verlauf eines in sich abgeschlossenen Einzelvorganges festlegen und diesem ein bestimmtes Merkmal zuordnen. Es muß den Einzelvorgang so definieren, daß er unendlich oft wiederholbar gedacht werden kann. Die Gesamtheit der Zahlen, die als Merkmale innerhalb einer gegebenen Einzelfolge möglich sind, heißt der Merkmalbereich. Das Merkmal selbst kann ein- oder mehrdimensional sein. Die definierte Gesamtheit umfaßt auch solche Folgen, die einen gesetzmäßigen mathematischen Aufbau in ihrer Anordnung besitzen. (Beispiel: Im Dezimalbruch $\sqrt{2}$ setze man an die Stelle jeder graden Ziffer eine „0" und jeder ungraden eine „1". Hier ist das Ergebnis eine Alternativfolge in bestimmter Anordnung.) Diese Definition bezieht sich auf die Erzeugungsart der Elementenfolgen und heiße daher „Erzeugungsprinzip". Aus der so bestimmten Gesamtheit wird durch eine zweite grundsätzliche Forderung, das „Konvergenzprinzip", eine Auslese getroffen. Sie lautet: Definiert A eine Teilmenge aus dem Merkmalbereich derart, daß man von jedem Element der Folge einwandfrei entscheiden kann, ob sein Element zu A gehört oder nicht, so soll n_A die Anzahl der Elemente mit dem Merkmal A unter den n ersten Elementen der Folge bedeuten. Bildet man dann

$$\frac{n_A}{n},$$

so soll dieser Quotient, die Größe der relativen Häufigkeit für das Vorkommen eines Elementes aus A, einem bestimmten Grenzwert zustreben, falls $n \to \infty$ gedacht wird:

$$\lim \frac{n_A}{n} = W_A, \; n \to \infty .$$

Aus der Gesamtheit solcher Folgen, die beiden Forderungen genügen, trifft v. M i s e s eine weitere Auslese, die sich auf die Anordnung ihrer Elemente bezieht. Dieses zweite Auslese-prinzip, das „Regellosigkeitsprinzip" lautet: Man hebe aus dem ganzen Merkmalbereich zwei einander fremde Teilmengen A und B heraus, für welche W_A und W_B nicht beide Null sind, und lasse alle Elemente unbeachtet, deren Merkmal weder zu A noch zu B gehört. Trifft man unter der Gesamtheit der herausgehobenen Elemente eine Stellenauswahl, so müssen innerhalb der so entstandenen Teilfolgen wieder die Grenzwerte:

$$\lim \frac{n'_A}{n'} = W'_A, \; \lim \frac{n'_B}{n'} = W'_B, \; n' \to \infty$$

existieren, und es muß gelten: $W'_A : W'_B = W_A : W_B$. Unter „Stellenauswahl" ist folgendes verstanden: Aus einer unendlichen Folge von Elementen wird durch Stellenauswahl eine Teilfolge gebildet, indem man eine Vorschrift angibt, durch die über die Zugehörigkeit oder Nichtzugehörigkeit des n-ten Elementes zur Teilfolge unabhängig von dem Merkmal dieses Elementes und höchstens unter Benutzung der Kenntnis der vorhergegangenen Elemente entschieden wird. Die Gesamtheit der Elementenfolgen, die allen drei Prinzipien genügen, ist identisch mit der Gesamtheit der Kollektive. W_A heißt die Wahrscheinlichkeit für das Vorkommen der zu A gehörenden Merkmale innerhalb des betreffenden Kollektivs.

Wir beschränken uns in der kritischen Prüfung auf die Frage, ob das Regellosigkeitsaxiom mit dem Konvergenzprinzip verträglich ist, da, wie in der Vorbemerkung gezeigt ist, dies die Kernfrage der v. M i s e s'schen Wahrscheinlichkeitstheorie darstellt. Zu dem Zwecke sei die Theorie der mathematischen reellen Zahlenfolgen angeführt. Hier ist jede Einzelfolge durch ihr Bildungsgesetz charakterisiert, das eine eindeutige Zuordnung aller positiven ganzen Zahlen zu einem Glied der Folge angibt, sodaß man zu jeder beliebig vorgegebenen ganzen Zahl das zugehörige Glied berechnen kann. Aus der Gesamtheit dieser Folgen werden die konvergenten nach folgendem Prinzip ausgesondert: Eine Zahlenfolge heißt konvergent, wenn ihr Bil-

dungsgesetz die mathematische Eigenschaft besitzt, daß sich zu jeder beliebig klein vorgegebenen Zahl $\varepsilon > 0$ ein N berechnen läßt, sodaß:

$$| A_{N+m} - A_N | < \varepsilon, \; m = 1, 2, \ldots .$$

Vergleicht man dieses Ausleseprinzip mit dem von v. M i s e s definierten, so ist zu beachten, daß in der Gesamtheit der zum Erzeugungsprinzip gehörenden Merkmalfolgen die mit mathematischem Bildungsgesetz nur eine Teilmenge bilden, und im allgemeinen grundsätzlich keine mathematische Zuordnung gegeben ist, aus der irgendetwas über die Größe der relativen Häufigkeit A_n in Abhängigkeit von n erschlossen werden könnte, wenn man von endlich vielen Elementen absieht. Mithin ist das in der Mathematik definierte Konvergenzprinzip als Ausleseprinzip auf die Gesamtheit der zum Erzeugungsprinzip gehörenden Folgen nicht anwendbar. Denn ein Prinzip, das vorschreibt, vorhanden gedachte Bildungsgesetze auf eine bestimmte mathematische Eigenschaft hin zu untersuchen, kann nicht auf eine Gesamtheit von Folgen angewandt werden, deren Bildungsgesetze als nicht bekannt vorausgesetzt werden. Das Regellosigkeitsprinzip verbietet aber die Annahme eines Bildungsgesetzes für die Elementenfolgen. Da jedoch die Existenz eines Bildungsgesetzes die Voraussetzung für die Annahme des Konvergenzprinzips bildet, definieren die zweite und dritte Forderung widerspruchsvolle Begriffe, falls man unter Konvergenz den in der Mathematik definierten Begriff versteht. Sollte v. M i s e s einen anderen Konvergenzbegriff meinen, so müßte dessen Definition mitgeteilt werden.

Aus diesen Überlegungen ergibt sich, daß man entweder bei Beibehaltung des infiniten Charakters der Folgen das Regellosigkeitsprinzip abändern oder finite Folgen einführen muß.

§ 2.

Definition eines Kollektivs aus endlich vielen Elementen und Durchführung der Grundoperationen.

Um ein Kollektiv aus endlich vielen Elementen zu definieren, denken wir uns endlich viele Merkmale gegeben, die insgesamt einen Merkmalbereich: M (A, B, . . .) bilden. Aus allen Merkmalen dieses Merkmalbereiches werde eine Folge von Elementen gebildet, indem man die Merkmale in einer Reihe in beliebiger Anordnung hinschreibt. Eine Folge aus n solchen Elementen werde mit E_n bezeichnet. Ist A eine Teilmenge des Merkmalbereiches M und n_A die Anzahl der insgesamt in E_n vorkommenden Elemente mit dem Merkmal A, so heißt:

$$A_n = \frac{n_A}{n}$$

die relative Häufigkeit des Vorkommens aller Elemente mit dem Merkmal A in bezug auf E_n. Es werde vorausgesetzt, E_n umfasse nur endlich viele Elemente, wobei n natürlich beliebig groß gewählt werden kann. Würde man auf ein E_n die vier v. Mises'schen Grundoperationen anwenden, so wäre man immer grundsätzlich in der Lage, die relative Häufigkeit im abgeleiteten E_n zu bestimmen, wenn auch die praktische Durchführung umständlich und langwierig wäre. Diese Tatsache hat wohl v. Mises zu der Behauptung veranlaßt (S. Erkenntnis 1930 Bd. I S. 279), daß eine Theorie endlicher Wahrscheinlichkeitsfolgen unmöglich wäre. Wir treffen aber aus der Gesamtheit der E_n eine Auslese nach folgendem Prinzip: Jedem Merkmal A, B, . . . aus M ordnen wir eine und nur eine reelle positive Zahl V_A, V_B, . . . zu. V soll im Bereich 0 bis 1 liegen und die Summe aller V soll gleich 1 sein. Die Gesamtheit T dieser Zahlen heißt „eine dem Merkmalbereich M zugeordnete Häufigkeitsverteilung" oder kurz „eine Verteilung zu M" und werde mit T (V_A, V_B, . . .) bezeichnet. Dann betrachten wir nur solche E_n, für welche die relative Häufigkeit aller Elemente des Merkmals A aus M um weniger als $\varepsilon_A > 0$ von der dem Merk-

mal A zugeordneten Zahl v_A abweicht. Die Summe $\varepsilon_A + \varepsilon_B + \varepsilon_C + \cdots$ in einer Folge E_n sei kleiner als $\varepsilon > 0$, wobei ε beliebig klein vorgeschrieben werden kann. Offenbar gibt es zu einem definierten M, T, ε beliebig viele E_n, welche diesen Anforderungen genügen. Ein solches E_n heiße im folgenden ein „Kollektiv" und die Gesamtheit aller E_n, die zu festen M, T, ε gehören, eine „Kollektivklasse". Eine Kollektivklasse liegt also vor, wenn folgende Bedingungen erfüllt sind:

1. Es muß ein Merkmalbereich M definiert sein.
2. M muß eine Häufigkeitsverteilung T zugeordnet sein.
3. Eine ε-Ungleichung soll erfüllt sein.

Beispiel:

(I) 1) M (A, B, C)

2) $T(V_A, V_B, V_C)$, $V_A + V_B + V_C = 1$

3) $V_A - \varepsilon_A < A_n < V_A + \varepsilon_A$,
$V_B - \varepsilon_B < B_n < V_B + \varepsilon_B$
$V_C - \varepsilon_C < C_n < V_C + \varepsilon_C$.

Dies gilt für alle $\varepsilon_A, \varepsilon_B, \varepsilon_C$, für die gilt: $\varepsilon_A + \varepsilon_B + \varepsilon_C < \varepsilon$.

Nachdem wir so den Begriff einer Kollektivklasse festgelegt haben, gehen wir zur Definition der Grundoperationen für diese Kollektivklassen über.

a) Auswahl: Ob eine durch Elementenauswahl aus einem E_n gebildete Folge E_m wieder zur gleichen Kollektivklasse gehört, ist im Einzelfalle grundsätzlich in endlich vielen Schritten mathematisch entscheidbar. In der Praxis kann man annehmen, E_m gehöre zur gleichen Kollektivklasse, und am Resultat einer Rechnung, das unter dieser Voraussetzung erhalten ist, kann man gegebenenfalls prüfen, ob die getroffene Voraussetzung über E_m berechtigt war oder nicht. Bei Behandlung des Anwendungsproblems werden wir ein hinreichendes Kriterium für die Gültigkeit der Auswahloperation ableiten.

b) Mischung: Wir besprechen sie zunächst am Beispiel (I) und fassen hier die Merkmale A und B zu einem einzigen D zusammen: A + B = D. Ist die bleibende Folge aus den Elementen der Merkmale D und C ein Kollektiv, und wie ist seine Verteilung? Die Antwort ergibt sich sofort. Der Merkmalbereich ist: M (D, C) und die Verteilung $T(V_D, V_C)$, wo:

$$V_D = V_A + V_B \quad , \quad V_D + V_C = 1 \, ;$$

es gilt die Ungleichung:

$$V_D - \varepsilon_A - \varepsilon_B < \frac{n_D}{n} < V_D + \varepsilon_A + \varepsilon_B ,$$

$$V_C - \varepsilon_C < \frac{n_e}{n} < V_C + \varepsilon_C \cdot$$

$$\varepsilon_A + \varepsilon_B = \varepsilon' \cdot$$

Daraus folgt, daß auch für die abgeleitete Folge die -Ungleichung erfüllt ist, wo $\varepsilon' + \varepsilon_c < \varepsilon$, d. h. es ist ein Kollektiv entstanden. Allgemein gilt folgendes: Faßt man in M mehrere Merkmale zu einem neuen zusammen und ordnet in E_n die entsprechenden Elemente dem neuen Merkmal zu, so besitzt das neue Merkmal als zugehörigen Verteilungswert die Summe der Verteilungswerte aller der Merkmale, aus denen es zusammengesetzt ist. M, T, und ε-Ungleichung sind vorhanden.

c) Teilung: Man denkt sich in E_n, das zu (I) gehört, alle Elemente des Merkmals C unbeachtet und alles übrige unverändert gelassen. Ist die dann übrigbleibende Elementenfolge mit M (A, B) ein Kollektiv und wie ist seine Verteilung? Nach Definition ist:

$$\frac{n_A}{n_A + n_B} = \frac{n_A/n}{n_A/n + n_B/n} = \frac{A_n}{A_n + B_n} = A'_n,$$

$$\frac{n_B}{n_A + n_B} = \frac{B_n}{A_n + B_n} = B'_n, \quad ;$$

A'_n und B'_n sind die gesuchten relativen Häufigkeiten, und da nach Voraussetzung insbesondere $\varepsilon_A + \varepsilon_B + \varepsilon_C < \varepsilon$ ist, gilt die Ungleichung:

$$\frac{V_A - \varepsilon}{(V_A + V_B) + \varepsilon} < A'_n < \frac{V_A + \varepsilon}{(V_A + V_B) - \varepsilon}$$

$$\frac{V_B - \varepsilon}{(V_A + V_B) + \varepsilon} < B'_n < \frac{V_B + \varepsilon}{(V_A + V_B) - \varepsilon}$$

Daraus folgt, die Summe ε' der auftretenden Abweichungen der Werte A'_n und B'_n von:

$$\frac{V_A}{V_A + V_B} \; ; \; \frac{V_B}{V_A + V_B}$$

läßt sich beliebig klein vorschreiben, da ε im Ausgangskollek-

tiv nach dieser vorgeschriebenen Summe der Abweichungen berechnet werden kann, d. h. die neue Elementenfolge hat den Merkmalbereich M (A, B), die Häufigkeitsverteilung:

$$T\left(\frac{V_A}{V_A+V_B}, \frac{V_B}{V_A+V_B}\right)$$

und erfüllt eine ε'-Ungleichung. Also ist sie ein Kollektiv. Allgemein gilt: Streicht man in M eine Anzahl Merkmale fort und in allen zugehörigen E_n die Elemente, die zu den in M gestrichenen Merkmalen gehören, so erhält man eine Kollektivklasse. Der Merkmalbereich ist die Gesamtheit der in M stehen gebliebenen Merkmale; der Häufigkeitsverteilungswert eines Merkmals ist gleich einem Quotienten, in dessen Nenner die Summe aller Verteilungswerte für die stehen gebliebenen Merkmale und in dessen Zähler der Verteilungswert des betreffenden Merkmals steht. Eine ε'-Ungleichung ist erfüllt. Also gehören die durch Teilung gebildeten Elementenfolgen zu einer Kollektivklasse mit bestimmtem M, T, und ε'.

d) Verbindung: Es seien zwei Kollektivklassen gegeben:

(I) $M_1\ (A, B); T_1 (V_A, V_B);$
$V_A-\varepsilon_A < A_n < V_A+\varepsilon_A\ ,\ V_B-\varepsilon_B < B_n < V_B+\varepsilon_B$

(II) $M_2(C,D); T_2(V_C, V_D);$
$V_C-\varepsilon_C < C_n < V_C+\varepsilon_C,\ V_D-\varepsilon_D < D_n < V_D+\varepsilon_D;$
$\varepsilon_A+\varepsilon_B < \varepsilon$
$\varepsilon_C+\varepsilon_D < \varepsilon$

Man denkt sich für gegebenes n je ein E_n aus (I) und (II) und setzt beide gliedweise in ihrer ursprünglichen Anordnung untereinander. Dann faßt man je zwei untereinander stehende Elemente als ein neues Element zusammen. Die neue Elementenfolge hat den Merkmalbereich: M (AC, AD, BC, BD). Es ist:

$$\frac{n_{AC}}{n}+\frac{n_{AD}}{n}+\frac{n_{BC}}{n}+\frac{n_{BD}}{n}=1.$$

Wenn auch (I) und (II) für die relativen Häufigkeiten gelten, so können bei der Verbindung Häufigkeitsverteilungen entstehen, die in ihren Einzelwerten stark voneinander abweichen. Man kann z. B. vorschreiben:

$$V_A = V_B = V_C = V_D = \frac{1}{2}$$

und 2 solche Folgen herausgreifen, in denen unter A nur einmal ein C, sonst immer D und unter einem B nur einmal ein D und sonst immer ein C steht. Dann ist:

$$\frac{n_{AC}}{n} = \frac{1}{n}, \quad \frac{n_{BC}}{n} = \frac{1}{n}.$$

Nun gilt allgemein:

$$\frac{n_{AC}}{n_A} + \frac{n_{AD}}{n_A} = 1, \quad \frac{n_{BC}}{n_B} + \frac{n_{BD}}{n_B} = 1,$$

$$\frac{n_{AC}}{n_A} \cdot \frac{n_A}{n} + \frac{n_{AD}}{n_A} \cdot \frac{n_A}{n} + \frac{n_{BC}}{n_B} \cdot \frac{n_B}{n} + \frac{n_{BD}}{n_B} \cdot \frac{n_B}{n} = 1.$$

Daraus folgt: Unter den durch Verbindung entstandenen E_{nn} gibt es solche, in denen:

$$\frac{n_{AC}}{n_A} = \frac{n_C}{n}, \quad \frac{n_{AD}}{n_A} = \frac{n_D}{n}, \quad \frac{n_{BC}}{n_B} = \frac{n_C}{n}, \quad \frac{n_{BD}}{n_B} = \frac{n_D}{n}.$$

Also entstehen unter anderen beliebig viele E_{nn}, deren relative Häufigkeit folgenden Ungleichungen genügen:

$$(V_A - \varepsilon)(V_C - \varepsilon) < \frac{n_{AC}}{n} = \frac{n_C}{n} \cdot \frac{n_A}{n} < (V_A + \varepsilon)(V_C + \varepsilon)$$

$$(V_D - \varepsilon)(V_A - \varepsilon) < \frac{n_{AD}}{n} = \frac{n_D}{n} \cdot \frac{n_A}{n} < (V_D + \varepsilon)(V_A + \varepsilon)$$

$$(V_C - \varepsilon)(V_B - \varepsilon) < \frac{n_{BC}}{n} = \frac{n_C}{n} \cdot \frac{n_B}{n} < (V_C + \varepsilon)(V_B + \varepsilon)$$

$$(V_D - \varepsilon)(V_B - \varepsilon) < \frac{n_{BD}}{n} = \frac{n_D}{n} \cdot \frac{n_B}{n} < (V_D + \varepsilon)(V_B + \varepsilon)$$

Daraus folgt, gibt man die Summe ε' der Abweichungen der Werte $n_{AC} : n$ von $V_{AC} = V_A \cdot V_C$ usw. vor, so läßt sich ε im Ausgangskollektiv berechnen, d. h. diese E_{nn} sind in einer Kollektivklasse enthalten, für welche:

(III) $$T\,(V_{AC}, V_{AD}, V_{BC}, V_{BD})$$

$$V_{AC} = V_A \cdot V_C, \qquad V_{AD} = V_A \cdot V_D,$$
$$V_{BC} = V_B \cdot V_C, \qquad V_{BD} = V_B \cdot V_D,$$

$$V_{AC} + V_{AD} + V_{BC} + V_{BD} = (V_A + V_B)(V_C + V_D) = 1,$$

und eine ε'-Ungleichung vorgelegt ist. Man erhält die neue Kollektivklasse, indem man die Häufigkeitsverteilung und ebenso die Merkmalbereiche rein kombinatorisch multipliziert und dann ihren Merkmalen entsprechend die Zuordnung vornimmt, wie die Formeln (III) angeben. Allgemein gilt folgendes: Bei der Verbindung zweier Kollektivklassen entsteht eine Gesamtheit neuer Elementenfolgen mit zweidimensionalem Merkmalbereich. Eine Teilmenge aus ihr gehört zu einer bestimmten Kollektivklasse, deren Verteilung und Merkmalbereich man dadurch erhält, daß man die Merkmalbereiche und die Häufigkeitsverteilungen je kombinatorisch multipliziert und ihrem Aufbau entsprechend die entstandenen Verteilungswerte den neuen Merkmalen zuordnet. In der zweiten Arbeit wird gezeigt, wann diese Teilmenge fast die Gesamtheit aller möglichen Verbindungen ausmacht. Zusammenfassend läßt sich also feststellen, auf die definierten Kollektive lassen sich die von v. Mises angegebenen Grundoperationen sinngemäß anwenden und die erhaltenen Resultate stimmen mit den von v. Mises angeführten überein. Auf den von v. Mises definierten Begriff „verbindbare“ Kollektive soll bei Behandlung des Anwendungsproblems näher eingegangen werden.

§ 3.

Der erste Fundamentalsatz mit seinen Sonderfällen.

Wir wollen nun untersuchen, ob und in welchem Sinne die von v. Mises formulierten Fundamentalsätze für die Theorie endlicher Kollektive gültig sind. Bei der Aufstellung einer Kollektivklasse muß man eine Häufigkeitsverteilung T vorgeben, d. h. ν positive Zahlen so bestimmen, daß ihre Summe 1 ist. Faßt man diese Zahlen als Koeffizienten eines Polynoms einer Veränderlichen t auf, so kann man auch sagen: Man muß ein Polynom in t mit positiven Koeffizienten aufbauen, das für $t = 1$ den Wert 1 annimmt:

$$f(t) = \sum_{x=o}^{x=\nu} v(x)\, t^x, \quad f(1) = 1.$$

Für das folgende erweist es sich zweckmäßig, die Merkmale des Bereiches M mit positiven ganzen Zahlen x zu bezeichnen. Der dem Merkmal x zugeordnete Verteilungswert soll v (x) heißen, und man faßt ihn als den Faktor von t^x auf. Man denkt sich m solcher Kollektivklassen mit ihren erzeugenden Funktionen:

$$f_k(t) = \sum_{x=o}^{x=\nu_k} v_k(x)\, t^x, \quad k = 1, 2, \cdots m\,.$$

gegeben und weiterhin für ein willkürliches, aber festes n eine m-fache Verbindung von je einem E_n aus jeder Kollektivklasse gebildet. Dabei kann man sich die Elementenfolgen untereinander gesetzt denken, sodaß man ein Rechtecksschema aus m wagerechten Reihen und n senkrechten Spalten erhält. In der Gesamtheit aller derartigen Matrizen wird eine Teilmenge enthalten sein, die zu der früher definierten Verbindungsklasse gehört. Diese erhält man, wenn man den Merkmalbereich und die Häufigkeitsverteilung der m-fachen Kollektivklasse nach den Regeln der kombinatorischen Multiplikation aus denen der Ausgangskollektive bildet und eine ε'-Ungleichung für die neue Verteilung vorschreibt. Dann kann man die ε-Abweichung in der Verteilung der Ausgangskollektive beweisen. Wir betrachten im folgenden nur Matrizen dieser Teilmenge und fragen: Wie groß

ist unter den n Spalten der Häufigkeitsverteilungswert für das Vorkommen solcher Spalten, in denen die Summe der m Merkmalzahlen einen Wert x besitzt? Wir suchen also ein neues Kollektiv zu bilden. Elemente des Kollektivs sind die Spalten der Matrix. Der Merkmalbereich besteht aus allen den positiven ganzen Zahlen x, die als Summe der m Elemente einer Spalte auftreten können. Die zugehörige Häufigkeitsverteilung findet man, indem man nach der Verbindungsregel alle $f_k(t)$ miteinander multipliziert und das neue Polynom unter Anwendung der Mischungsregel wieder nach Potenzen von t ordnet. Die Koeffizienten mögen $w_m(x)$ genannt sein; diese Faktoren der entstandenen Potenzen von t bilden die gesuchte Verteilung T. Die Summe aller Faktoren $w_m(x)$ ist 1, da das Produkt aller $f_k(t)$ für $t = 1$ den Wert 1 besitzt. Wir bringen zunächst im folgenden die mathematischen Ableitungen, welche sich auf die Abschätzung des Verteilungswertes $w_m(x)$ für große Werte von m beziehen, d. h. den ersten Fundamentalsatz, und zeigen darauf, daß eine ε'-Ungleichung erfüllbar ist. Dabei schließen wir uns eng an die Ableitungen an, die v. Mises in Bd. 1 § 8, 2—4 durchführt. Der erste Fundamentalsatz besagt, daß:

$$w_m(x) \to \text{const.}\ e^{-u^2},\ m \to \infty,$$

wobei die Abhängigkeit u (x) noch näher angegeben wird. Um den Beweis zu führen, denken wir uns im Anschluß an v. Mises in den erzeugenden Funktionen $f_k(t)$ die Unabhängige $t = e^{iz}$ eingesetzt und erhalten:

$$f_k(t) = \sum_{x=0}^{x=\nu_k} v_k(x)\, e^{xzi}$$

Dann bilden wir:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} f_1(z) \cdot f_2(z) \cdot \cdot \cdot f_m(z)\, e^{-xzi}\, dz$$

und zeigen, daß dies Integral den Wert $w_m(x)$ besitzt. Dies ist leicht einzusehen, wenn man sich das Produkt der $f(z)$ ausgerechnet und nach Potenzen von e^{iz} geordnet denkt. Nach unseren Voraussetzungen erhält man dann:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} [w_m(0 + w_m(1)\, e^{zi} + \cdot \cdot + w_m(x)\, e^{xzi} + \cdot \cdot]\, e^{-xzi}\, dz.$$

Bei gliedweisem Integrieren werden dann alle Integrale Null, bis auf dasjenige, wo im Integranden:

$$w_m(x) \, e^{xzi} \cdot e^{-xzi} = w_m(x)$$

erscheint, und dies Integral hat den Wert $w_m(x)$. Also gilt: (Bd. I § 8 Gl. 27)

$$w_m(x) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} f_1(z) \cdot f_2(z) \cdot \cdot \cdot f_m(z) \cdot e^{-xzi} \, dz$$

Wir können nun die von v. Mises im Bd. I § 2 Gl. 6 und 8 angegebenen Ausdrücke für Mittelwert und Streuung einer Verteilung ohne weiteres übernehmen. Dann wird für k = 1, 2, 3, . . . m:

$$a_k = \sum_{x=0}^{x=\nu_k} x \cdot v_k(x),$$

$$r_k^2 = \sum_{x=0}^{x=\nu_k} (x - a_k)^2 \, v_k(x).$$

In Bd. I § 8, 4 beweist v. Mises, daß das letztgenannte Integral durch folgende Formel abgeschätzt wird: (§ 8 Gl. 33)

$$\text{(IV)} \qquad \lim s_m \sqrt{2} \; w_m(x) = \frac{1}{\sqrt{\pi}} e^{-u^2}, \qquad m \to \infty,$$

gleichmäßig für alle $|u| < U$, wo:

$$x = b_m + s_m \sqrt{2} \cdot u$$
$$b_m = a_1 + a_2 + \cdot \cdot + a_k + \cdot \cdot + a_m$$
$$s_m^2 = r_1^2 + r_2^2 + \cdot \cdot + r_k^2 + \cdot \cdot + r_m^2.$$

Die Bedingungen, unter denen (IV) gilt, sind bei v. Mises Bd. I S. 212 zusammengestellt. Damit ist bewiesen, daß wir den von v. Mises S. 212 ausgesprochenen ersten Fundamentalsatz, der durch Gleichung (IV) ausgedrückt wird, wörtlich übernehmen können. Er lautet:

„Die Summenbildung $w_m(x)$ aus m verbindbaren Kollektiven mit arithmetischen Verteilungen führt bei großem m auf das Gauß'sche Gesetz."

Aus dieser an sich rein mathematischen Tatsache wollen wir nun eine Folgerung über die Bauart unserer durch Verbindung erhaltenen Matrix ableiten. Bedeutet genau wie bei v. Mises S. 211 $W_m(x)$ die Summenfunktion zu $w_m(x)$, d. h. die Summe aller $w_m(\xi)$, für welche $\xi \leq x$, und wird $W_m(x_2) - W_m(x_1) = W_m(x_2, x_1)$ gesetzt, so gilt für

$x_2 = b_m + \zeta$ und $x_1 = b_m - \zeta$:

$$W_m(b_m + \zeta,\ b_m - \zeta) \sim \frac{1}{\sqrt{\pi}} \int_{-u'}^{o} e^{-u^2} du + \frac{1}{\sqrt{\pi}} \int_{o}^{u'} e^{-u^2} du,$$

$$u' = \frac{\zeta}{s_m \sqrt{2}}.$$

Setzt man:

$$\Theta(u') = \frac{2}{\sqrt{\pi}} \int_{o}^{u'} e^{-u^2} du,$$

so ist:

$$W_m(b_m + \zeta, b_m - \zeta) \sim \Theta\ u') = \Theta\left(\frac{\zeta}{\sqrt{2}\ s_m}\right),$$

wo:

$$\Theta(x) = \frac{2}{\sqrt{\pi}} \int_{o}^{x} e^{-y^2} dy$$

das bekannte Gauß'sche Integral ist. Wir formen weiter um, indem wir bilden:

$$b'_m = \frac{a_1 + a_2 + \cdots + a_m}{m}$$

$$s'^2_m = \frac{r_1^2 + r_2^2 + \cdots + r_m^2}{m^2}$$

und $\zeta = mz$ einsetzen. Dann wird, da:

$$W_m(b_m + \zeta,\ b_m - \zeta) = W'_m(b'_m + z,\ b'_m - z),$$

(V) $$W'_m(b'_m + z,\ b'_m - z) \sim \Theta\left(\frac{z}{s'_m \sqrt{2}}\right).$$

Hier ist zu beachten, daß s_m wie $\sqrt{m}$ nach der in Bd. I S. 201 angegebenen Voraussetzung unendlich groß wird, also s'_m wie $m^{-1/2}$ unendlich klein wird und $\Theta(x) \to 1$ wenn $x \to \infty$. Die Formel (V) stellt den mathematischen Ausdruck für das erste Gesetz der großen Zahlen dar. Damit sind wir in der Lage, dieses Gesetz, das v. Mises Bd. I S. 180 formuliert, als rein mathematischen Anordnungssatz für die durch m-fache Verbindung erhaltene Matrix auszusprechen, falls wir das Erfülltsein einer ε'-Ungleichung als bewiesen vorwegnehmen. Es lautet:

„In der durch m-fache Verbindung erhaltenen Matrix kann man die relative Häufigkeit der Spalten, in denen das arithmetische Mittel der in ihr stehenden Elemente um weniger als z von dem Erwartungswert abweicht, beliebig nahe an den Wert 1 heranbringen, wenn nur m genügend groß gewählt wird."

Kürzer und weniger genau kann man sagen: Das arithmetische Mittel fast aller Spalten der Matrix ist nahezu gleich dem Erwartungswert, wenn nur m genügend groß ist. Gehören insbesondere alle m Kollektive derselben Kollektivklasse an, so ist der Erwartungswert der m-fachen Verbindung $b'_m = a$, d. h. gleich dem Mittelwert a der Verteilung der betreffenden Kollektivklasse. Ist darüber hinaus diese Kollektivklasse alternierend, d. h.

$$f_k(t) = v_k(o) + v_k(1)\,t = p + q\,t$$

so ist: $b_m = ma_1 = ma$ und $a = q$, d. h. $b'_m = q$. Ferner ist: $s_m^2 = r_1^2 + + \cdots + r_m^2 = mr_1^2 = mr^2$, $r^2 = pq$, also: $s_m^2 = mpq$ und $s'^2_m = \frac{pq}{m}$, $s'_m = \sqrt{\frac{pq}{m}}$. Daraus folgt, das Bernoullische Theorem:

$$W_m(q + z, q - z) \sim \Phi\left(z\sqrt{\frac{m}{pq}}\right).$$

Denkt man sich also m Kollektive, die zu einem Alternativbereich M (A, B) und der zugeordneten Alternativverteilung T (p, q) gehören, durch Verbindung zusammengefaßt, so entsteht eine Matrix aus m Reihen und n Spalten. Die relative Häufigkeit solcher Spalten, in denen die relative Häufigkeit des Vorkommens aller zum Merkmal B gehörenden Elemente um weniger als z von q abweicht, weicht beliebig wenig vom Wert 1 ab, wenn nur m genügend groß genommen wird. Aus dem Bernoullischen Theorem ziehen wir eine wichtige Folgerung: Da jede Kollektivklasse durch Mischung in eine alternierende Kollektivklasse verwandelt werden kann, so besagt das Bernoullische Theorem auch: Besteht die Matrix aus m Kollektiven derselben Kollektivklasse, so treten die zum Merkmal A gehörigen Elemente in fast allen Spalten nahezu mit der gleichen relativen Häufigkeit auf, mit der sie je in den zur m-fachen Verbindung benutzten einzelnen Kollektiven vorkommen. Auf weitere Folgerungen aus dem Fundamentalsatz wird bei Behandlung des Anwendungsproblems näher eingegangen werden. Wir zeigen noch, man kann ε in den Ausgangskollektivklassen von vornherein so klein wählen, daß alle E_n^m, die

durch m-fache Verbindung erhalten werden, eine Häufigkeitsverteilung für die Summe x besitzen, die vorgeschrieben wenig von:

$$\text{(VI)} \qquad \frac{1}{s_m \sqrt{2\pi}} \, e^{-\frac{(x-b_m)^2}{2 s_m^2}}$$

abweicht, und zwar gleichmäßig für alle $|u| < U$, wo:

$$u = (x - b_m) \cdot s_m \sqrt{2}.$$

Diese Wahl von ε ist grundsätzlich immer möglich. η sei als Abweichung vom Ausdruck (VI) vorgegeben. Man teilt η auf: $\eta_1 + \eta_2 < \eta$ und bestimmt nach der v. Mises'schen Abschätzung im Bd. I § 8, 2 derart m, daß für ein $m > m_1$ $w_m(x)$ um weniger als η_1 von (VI) abweicht, und zwar für alle x gleichmäßig, deren zugehöriges $|u| < U$ ist. Darauf bestimmt man ε zu jedem festgewählten $m > m_1$: Die Summenverteilung $w_m(x)$ ist eine Summe aus angebbar endlich vielen Produkten, die aus angebbar endlich vielen Faktoren $v_K(\xi)$ bestehen. Sie liegt zwischen zwei anderen Summen, die aus den gleichen Produkten wie $w_m(x)$ aufgebaut sind, in denen an Stelle von $v_K(\xi)$ der Ausdruck $v_K(\xi) \pm \varepsilon$ als Einzelfaktor steht. Da in der Summe alles bestimmt ist bis auf ε, kann man sie in ein Polynom nach Potenzen von ε entwickeln und darauf ε so klein wählen, daß die beiden Summen, die als Schranken für $w_m(x)$ auftreten, höchstens um η_2 von $w_m(x)$ abweichen, und zwar gleichmäßig für alle in Frage kommenden x, für welche $|u| < U$ ist. Damit ist die Gültigkeit einer ε'-Ungleichung allgemein nachgewiesen.

Wenn wir die Ableitungen über den ersten Fundamentalsatz zusammenfassen, so stellen wir fest, daß dieser in der Theorie endlicher Merkmalfolgen mit dem von v. Mises angegebenen Wortlaut gültig ist und daß seine Folgerungen bestimmte rein mathematische Aussagen über gewisse durch Verbindung entstandene Matrizen enthalten.

§ 4.

Das Bayes'sche Problem.

Um zu zeigen, daß auch der zweite Fundamentalsatz aus der v. M i s e s'schen Theorie in die Theorie endlicher Kollektive als rein mathematischer Anordnungssatz sich einfügt, besprechen wir das Bayes'sche Problem als einfachsten Sonderfall des allgemeinen Satzes. Dies geschieht, weil man hier den Grundgedanken in einfachster Form hervorheben kann, ohne allzu verwickelte Formulierungen zu benutzen. Wir denken uns einen Merkmalbereich M (A, B) und zu ihm gehörig ν Verteilungen $T_\mu(x_\mu, 1-x_\mu)$ $\mu = 1, 2, 3, \ldots \nu$. ε sei für alle Verteilungen gleich gewählt. Aus jeder der so definierten ν Kollektivklassen denken wir uns je m Kollektive, die der Verbindungsregel genügen, herausgegriffen und verbunden, sodaß man im ganzen ν Matrizen erhält und in jeder m Kollektive aus einer einzigen Kollektivklasse stehen. Die aus der Verteilung $T_\mu(x_\mu, 1-x_\mu)$ gebildete Matrix sei mit R_μ bezeichnet. Sie besteht aus n Spalten mit je m Elementen. Die relative Häufigkeit aller solchen Spalten in der Matrix R_μ, in denen das Element A genau $m_1 \leq m$ mal auftritt, ist nach der als gültig angenommenen Verbindungsregel durch den Verteilungswert:

$$\binom{m}{m_1} x_\mu^{m_1} (1-x_\mu)^{m-m_1}$$

abgeschätzt. Aus dem ersten Fundamentalsatz folgt für alternierende Kollektive die Formel 25 Bd. I § 5 bei v. M i s e s. Aus ihr geht hervor, daß bei einer m-fachen Verbindung von m E_n aus einer einzigen Kollektivklasse mit vorgeschriebenem M (A, B), T (V_A, V_B) und ε die relative Häufigkeit derjenigen Spalten, die das Element B genau m_1 mal enthalten, eine Nullfolge durchläuft, falls man aus der Kollektivklasse immer mehr Kollektive von hinreichend großer Elementenzahl verbindet, also m genügend groß werden läßt, d. h. bei großem m kommen in der Matrix R_μ in bezug auf die Gesamtheit ihrer Spalten nur sehr wenig Spalten vor, die das Element mit dem Merkmal A

genau m_1 mal enthalten. Wir werfen die Frage auf: In welchen Matrizen R_μ stehen die meisten aller überhaupt vorkommenden Spalten mit m_1 Elementen, die zum Merkmal A gehören? Zur Beantwortung dieser Frage denken wir uns aus der Gesamtheit der R_μ als Merkmalbereich ein Kollektiv E_n' gebildet. Als Verteilung nehmen wir hierfür eine Gleichverteilung an, sodaß jedes Merkmal R_μ den Verteilungswert $1:\nu$ erhält. (V. Mises legt allgemeiner eine beliebige Verteilung zu Grunde für die Merkmale R_μ. Das Endergebnis ändert sich dadurch nicht.) In E_n' setzen wir für jedes Element die betreffende Matrix ein und erhalten dadurch eine Gesamtmatrix. Nach der Verbindungs- und Teilungsregel ist die relative Häufigkeit aller Spalten mit m_1 Elementen, die zum Merkmal A gehören und zu Matrizen mit dem Merkmal R_μ im Verhältnis zu allen in der Gesamtmatrix vorkommenden Spalten mit m_1 Elementen des Merkmals A durch folgenden Verteilungswert abgeschätzt:

$$\text{(VII)} \qquad \frac{\binom{m}{m_1} x_\mu^{m_1} (1-x_\mu)^{m-m_1}}{\sum_{\mu=1}^{\mu=\nu} \binom{m}{m_1} x_\mu^{m_1} (1-x_\mu)^{m-m_1}}$$

(vergl. Bd. I S. 152). Man kann die Anzahl der Zahlenpaare $(x_\mu, 1-x_\mu)$ sich größer und größer gewählt denken. Damit wächst auch die Zahl der Matrizen. Die Ausdrücke (VII), die man dabei erhält, nähern sich dem Wert einer kontinuierlichen Funktion als Grenzwert an, falls $\nu \to \infty$. Dieser Wert ist für $x_\mu = x$:

$$\text{(VII a)} \qquad w_m(x) = \frac{x^{m_1}(1-x)^{m-m_1}}{\int_0^1 x^{m_1}(1-x)^{m-m_1}\,dx}$$

Sie heißt Verteilungsdichtefunktion für die Verteilung (VII). Von dieser Funktion läßt sich nach v. Mises Bd. I § 6, 2 zeigen, daß sie, wenn $m \to \infty$ und $\alpha = \frac{m_1}{m}$ fest ist, sich gleichmäßig folgender Funktion annähert:

$$\text{(VII b)} \qquad w_m(x) \sim \sqrt{\frac{m}{2\alpha(1-\alpha)\pi}}\, e^{-\frac{m(x-\alpha)^2}{2\alpha(1-\alpha)}}$$

Die Annäherungsbedingungen sind Bd. I § 6 S. 156 ff. angegeben. Bedeutet:

$$W_m(x) = \int_0^x w_m(\xi)\, d\xi,$$

so ist nach Bd. I § 6 Gl. 20 für:

$$x_1 = \alpha - z < \xi < \alpha + z = x_2$$

$$(VII c) \qquad W_m(\alpha+z,\ \alpha-z) \sim \Theta\left(z\sqrt{\frac{m}{2\alpha(1-\alpha)}}\right)$$

Wählt man m genügend groß, so strebt die rechte Seite gegen den Wert 1. Setzen wir die Gültigkeit einer ε'-Ungleichung voraus, so können wir folgendes über die Matrizen aussagen: Die Spalten, welche das Element A genau $m_1 = \alpha \cdot m$ mal enthalten, gehören nahezu sämtlich zu solchen Matrizen, deren zugehörige Verteilung folgender Ungleichung genügt:

$$\alpha - z < x_\mu < \alpha + z,$$

wenn nur m bei konstantem α genügend groß ist, d. h. die Spalten mit m_1 A-Elementen stehen fast in ihrer Gesamtheit in solchen Matrizen, deren zugehöriger Verteilungswert von α nur wenig abweicht. Damit ist das Bayes'sche Problem als rein mathematischer Anordnungssatz für endliche Kollektive erhalten, wobei die Beweise von v. M i s e s fast wörtlich übernommen sind.

Wir haben noch zu zeigen, daß man grundsätzlich in der Lage ist, in den Ausgangskollektiven ε so klein zu wählen, daß alle Kollektive, die den geforderten Operationen unterworfen werden, den für das Merkmal A verlangten Bedingungen genügen, d. h. es muß ε so klein gewählt werden können, daß folgende Ungleichung erfüllt wird:

$$(VIII) \qquad \left| \frac{(x_\mu \pm \varepsilon)^{m_1}\,[1-(x_\mu \pm \varepsilon)]^{m-m_1}}{\sum_{\mu=1}^{\nu} (x_\mu \pm \varepsilon)^{m_1}\,[1-(x_\mu \pm \varepsilon)]^{m-m_1}} - \sqrt{\frac{m}{2\alpha(1-\alpha)\pi}}\, e^{\frac{m\,(x-\alpha)^2}{2\alpha(1-\alpha)}} \right| < \eta$$

wo η beliebig klein und die Abschätzung für alle $\alpha - z < x_\mu < \alpha + z$ gleichmäßig sein soll. Man denkt sich η aufgeteilt: $\eta_1 + \eta_2 + \eta_3 < \eta$. Dann bestimmt man m so groß, daß gleichmäßig für alle x des Bereiches $\alpha - z < x < \alpha + z$ $w_m(x)$ um weniger als η_1 von der rechten Seite in (VIIb) abweicht. Darauf wählt man ν so groß, daß alle Ausdrücke (VII) für $\alpha - z < x < \alpha + z$ von dem Ausdruck (VIIa), in dem $x_\mu = x$ ist, gleichmäßig für alle $\alpha - z < x < \alpha + z$ um höchstens η_2 abweichen. Dies ist möglich für jedes festgewählte m. Nachdem so

über m, ν, x verfügt ist, kann man noch ε dahin abschätzen, daß der erste Summand in (VIII) um weniger als η_3 von dem Ausdruck (VII) abweicht, und zwar gleichmäßig für alle x des gegebenen Bereiches. Auch diese Abschätzung ist grundsätzlich möglich. Also ist (VIII) erfüllbar, und da die Abschätzung gleichmäßig für alle:

$$\alpha - z < x < \alpha + z$$

ist, gilt eine entsprechende Abschätzung für (VIIc).

Der zweite Fundamentalsatz von v. Mises läßt sich in ähnlichem Sinne als rein mathematischer Anordnungssatz in der Theorie endlicher Kollektive auffassen. Damit ist gezeigt, daß für die Theorie endlicher Kollektive die Fundamentalsätze gültig sind, wenn man die Verbindbarkeit der Kollektive voraussetzt. Diese Sätze erhalten jedoch erst ihre volle Bedeutung in der zweiten Arbeit, wo gezeigt wird, daß man bei Annahme einer einfachen Voraussetzung über diese Kollektive deren Verbindbarkeit „fast immer“ annehmen kann.

§ 5.

Kollektive mit geometrischer Verteilung.

Wir wollen kurz zeigen, daß die Theorie der endlichen Kollektive auch die von v. Mises definierten Kollektive mit geometrischer Verteilung umfaßt.

M sei ein Bereich der reellen X-Achse, der von $0 < x < \xi$ sich erstreckt, und w (x) eine Funktion derart, daß:

$$\int_0^{\xi} w(x)\, dx = 1.$$

w (x) sei reell und positiv und heißt „Verteilungsdichtefunktion". Ferner seien A, B, C Teilbereiche, die insgesamt M ausmachen, aber keinen Punkt gemeinsam haben. E_n sei eine Folge von Zahlen aus M und $n_A, n_B, n_C (n_A + n_B + n_C = n)$ die Anzahl derjenigen Zahlen, die in A, B, C liegen. Wir betrachten die Gesamtheit aller E_n, für welche:

$$W_A - \varepsilon_1 < \frac{n_A}{n} < W_A + \varepsilon_1,\ W_B - \varepsilon_2 < \frac{n_B}{n} < W_B + \varepsilon_2,\ W_C - \varepsilon_3 < \frac{n_C}{n} < W_C + \varepsilon_3,$$

$$\text{(IX)} \qquad W_A = \int_A w(x)\, dx, \quad W_B = \int_B w(x)\, dx, \quad W_C = \int_C w(x)\, dx,$$

$$W_A + W_B + W_C = 1,\ \varepsilon_1 + \varepsilon_2 + \varepsilon_3 < \varepsilon.$$

Allgemein lassen sich ν solcher Bereiche auf M angeben und somit Kollektivklassen mit einem Merkmalbereich aus ν Merkmalen bilden. Kollektivklassen, deren Merkmalbereich aus einem kontinuierlichen Bereich M und deren Verteilungswerte aus einem M zugehörigen Integral abgeleitet werden, seien kurz „Kollektive mit geometrischer Verteilung" genannt. Die Definition gilt sinngemäß auch für mehrdimensionale Gebilde. Es soll gezeigt werden, daß alle aus Kollektiven mit geometrischer Verteilung durch die Grundoperationen abgeleiteten Kollektive wieder geometrische Verteilungen besitzen. Bei der Mischung ist dies ohne weiteres klar. Die Richtigkeit der Teilung zeigen wir am Beispieltyp (IX). Gegeben sei E_n, das den Bedin-

gungen (IX) genügt. Wir lassen in E_n alle Zahlen unbeachtet, die in C liegen. Ist die übrigbleibende Merkmalfolge ein Kollektiv mit geometrischer Verteilung? Es ist:

$$\frac{n_A}{n_A+n_B}=\frac{A_n}{A_n+B_n}\,,\quad \frac{n_B}{n_A+n_B}=\frac{B_n}{A_n+B_n}\,,$$

$$\frac{W_A-\varepsilon}{(W_A+W_B)+\varepsilon}<\frac{A_n}{A_n+B_n}<\frac{W_A+\varepsilon}{(W_A+W_B)-\varepsilon}$$

$$\frac{W_B-\varepsilon}{(W_A+W_B)+\varepsilon}<\frac{B_n}{A_n+B_n}<\frac{W_B+\varepsilon}{(W_A+W_B)-\varepsilon}$$

d. h. der neue Merkmalbereich ist A + B und die Verteilungsdichtefunktion:

$$\frac{w(x)}{\int\limits_{A+B} w(x)\,dx}\,.$$

Also liegt ein Kollektiv mit geometrischer Verteilung vor. Allgemein gilt: Teilt man einen geometrischen Bereich Q, wo:

$$\int\limits_Q w(q)\,dq=1.$$

in eine endliche Anzahl von Teilbereichen auf und ordnet jedem dieser Teilbereiche A den Verteilungswert zu:

$$W_A=\int\limits_A w(q)\,dq<1,$$

so sei ε vorgegeben und ein E_n, das zu dieser Verteilung paßt, bestimmt. Streicht man im Merkmalbereich M einzelne Merkmale und die zu diesen Merkmalen gehörigen Elemente in E_n, so bleibt ein Kollektiv E'_n übrig, dessen Verteilungswert für ein Merkmal A in bezug auf den neuen Merkmalbereich gleich einem Quotienten ist, in dessen Nenner die Summe aller Verteilungswerte stehen, die den übriggebliebenen Merkmalen im alten Merkmalbereich zukamen, während im Zähler W_A auftritt. Die dieser Teilung zugrunde liegende Verteilungsdichtefunktion ist ein Quotient mit w (q) im Zähler und der bezeichneten Summe aus den übriggebliebenen Verteilungswerten im Nenner.

Bei der Operation der Verbindung denken wir uns auf der X-Achse einen Bereich M_1 $0\leqq x\leqq\xi$ gegeben und w (x), sodaß:

$$\int\limits^{\xi} w(x)\,dx=1\;;$$

auf der Y-Achse sei ein Bereich $M_2\ 0 \leqq y \leqq \eta$ gegeben, und es sei für $v(y)$:

$$\int_0^{\eta} v(y)\,dy = 1 .$$

Wir denken uns aus beiden geometrischen Bereichen und Verteilungsdichtefunktionen je eine Kollektivklasse mit vorgeschriebenem ε gebildet. In der XY-Ebene ist ein Rechteck mit den Seiten ξ . η gegeben. Durch die Bildung der Merkmalbereiche für die beiden Kollektivklassen wird dieses Rechteck bei Aufteilung der linearen Bereiche η , ξ in Teilrechtecke zerlegt.

Sind:

$$W_A = \int_A w(x)\,dx\ ,\ W_B = \int_B v(y)\,dy$$

die zu A bzw. B gehörenden Verteilungswerte, so gilt für verbindbare Kollektive

$$(W_A - \varepsilon)(W_B - \varepsilon) < \frac{n_A}{n} \cdot \frac{n_B}{n} < (W_A + \varepsilon)(W_B + \varepsilon) .$$

Also ist die zum Teilrechteck mit den Seiten A und B gehörige Verteilung:

$$W_A \cdot W_B = \int_A w(x)\,dx \cdot \int_B v(y)\,dy = \iint_{A\,B} w(x) \cdot v(y)\,dx\,dy .$$

Daraus ergibt sich als Verteilungsdichtefunktion: $w(x) \cdot v(y)$ für den Punkt xy im Gesamtrechteck, wobei:

$$\int_0^{\xi}\int_0^{\eta} w(x) \cdot v(y)\,dx\,dy = 1 .$$

Mithin entsteht durch die Verbindung wieder ein Kollektiv mit geometrischer Verteilung. Sind allgemein m Verteilungsdichtefunktionen $w_\mu(x_\mu)$, $\mu = 1, 2, \ldots m$ gegeben, wo:

$$\int_0^{\xi_\mu} w_\mu(x_\mu)\,dx_\mu = 1,$$

so entsteht durch m-fache Verbindung die Dichtefunktion:

$$w_1(x_1) \cdot w_2(x_2) \ldots w_m(x_m)$$

für einen Punkt im m-dimensionalen Raum. Das m-fache Integral über der Dichtefunktion im m-dimensionalen Raum hat den Wert 1.

Nachdem wir gezeigt haben, daß die Kollektive, die aus Kollektiven mit geometrischer Verteilung durch eine der Grundoperationen gebildet sind, wieder geometrische Verteilungen besitzen, besprechen wir noch den ersten Fundamentalsatz für geometrische Verteilungen. Gegeben seien m Verteilungsdichtefunktionen: $w_\mu(x_\mu)$, für die:

$$\int_{-\infty}^{+\infty} w_\mu(x_\mu)\, dx_\mu = 1 .$$

Bildet man zu jeder Funktion eine Kollektivklasse mit einem Merkmalbereich und mit der zugeordneten Verteilung aus je ν Elementen, so erhält man bei vorgegebenem ε m Kollektivklassen, bei deren Verbindung der erste Fundamentalsatz gilt, wenn die Kollektive verbindbar sind. Die dabei auftretende Summenverteilung ist aus einer Dichtefunktion ableitbar, und der Sinn des ersten Fundamentalsatzes für geometrische Verteilungen ist der Nachweis, daß diese stetige Summendichtefunktion in ihrem ganzen Verlauf gegen die Funktion:

$$\frac{1}{s_m \sqrt{2\pi}}\, e^{-\frac{(x-b_m)^2}{2 s_m^2}}$$

konvergiert.

Dabei ist die v. Mises'sche Definition des Mittelwertes und der Streuung für geometrische Verteilungen (Bd. I § 2, 2 Gln. 6 und 8) ohne Änderung übernommen. Insbesondere ist also:

$$a_\mu = \int_{-\infty}^{\infty} x_\mu \cdot w_\mu(x_\mu)\, dx_\mu , \; b_m = a_1 + \cdots a_\mu + \cdots + a_m ;$$

$$r_\mu^2 = \int_{-\infty}^{\infty} (x_\mu - a_\mu)^2 w_\mu(x_\mu)\, dx_\mu, \; s_m^2 = r_1^2 + r_2^2 + \cdots + r_\mu^2 + \cdots + r_m^2 .$$

Im Bd. I § 8, 5 ist der Beweis für die Richtigkeit dieser Abschätzung geführt, und dort sind auch die Bedingungen angegeben, unter denen sie gilt. Die Beweisführung selbst bleibt unverändert bestehen. Wir stellen damit fest, daß auch der von v. Mises bewiesene erste Fundamentalsatz für geometrische Verteilungen wortgetreu für die Theorie endlicher Kollektive Geltung hat.

LEBENSLAUF.

Am 2. Januar 1901 bin ich als Sohn des Maschinensetzers Wilhelm Blume in Paderborn geboren. Von Ostern 1907 bis Ostern 1913 besuchte ich die Volksschule in Paderborn und von Ostern 1913 an das Gymnasium Theodorianum in Paderborn, das ich Ostern 1922 nach Ablegung der Reifeprüfung verließ, um Mathematik und Naturwissenschaften zu studieren. Das Studium führte mich an die Universitäten in Göttingen, Berlin und Münster i. W. Am 12. Januar 1927 legte ich in Münster die wissenschaftliche Prüfung für das Lehramt an Höheren Schulen und nach zweijähriger Ausbildung als Studienreferendar am 20. März 1929 das Pädagogische Examen für das Lehramt an Höheren Schulen ab. Darauf war ich an den Höheren Schulen der Stadt Dortmund als Assessor beschäftigt und bin seit dem 1. Oktober 1930 Studienrat am Bismarck-Realgymnasium in Dortmund.